Boston Strategic Partners, Inc.

Practical Considerations to Big Data Analysis

"Lessons learned from electronic health records (EHR) analysis"

Victor Khangulov, PhD

Innovations in Design, Analysis, and Dissemination: Frontiers in Biostatistics Conference

Kansas-Western Missouri chapter of the ASA

and

Department of Biostatistics in the Kansas University Medical Center

April, 2015

Boston Strategic Partners, Inc.
4 Wellington St. Suite 3
Boston, MA 02118
www.bostonsp.com

Contents

1. Who we are, what we do, and the data involved
2. Learnings about patient / group selection
3. Learnings about clinical education / context of the project
4. Learnings about making strategic recommendations and leveraging data
5. Acknowledgements
6. Q&A time

Boston Strategic Partners is a healthcare consulting firm with three practice areas

Business and Clinical Strategy

- Life-Cycle Planning
- M&A Advisory
- Primary Market Research
- Forecasts
- Pricing and Reimbursement Plans

Health Economics

- Retrospective Studies
- Epidemiology/Burden of Disease
- Budget Impact Modeling
- Claims Modeling

Medical Communications

- Publication Strategy
- Scientific Research Publications
- Abstracts and Posters
- Meeting Facilitation

Health economics and outcomes research (HEOR) complements clinical development information (ie, efficacy, safety, quality) to guide decision makers regarding patient access to specific drugs, services and their associated costs.

Holtorf A-P, Brixner D, Bellows B, Keskinaslan A, Dye J, Oderda G. Current and Future Use of HEOR Data in Healthcare Decision-Making in the United States and in Emerging Markets. *Am Heal drug benefits*. 2012;5(7):428–38.

For our clients, we use EHR to support their products that are in-development as well as evaluation of currently marketed products

R&D / Medical Affairs

- Clinical trial design
- Better understanding of disease and care pathways
- Understanding how drugs are actually utilized in “real world” scenarios (including off-label use)

Health Economics

- Analyze budget impact to the hospital budget
- Determine cost effectiveness of medications
- Improve coverage and payments

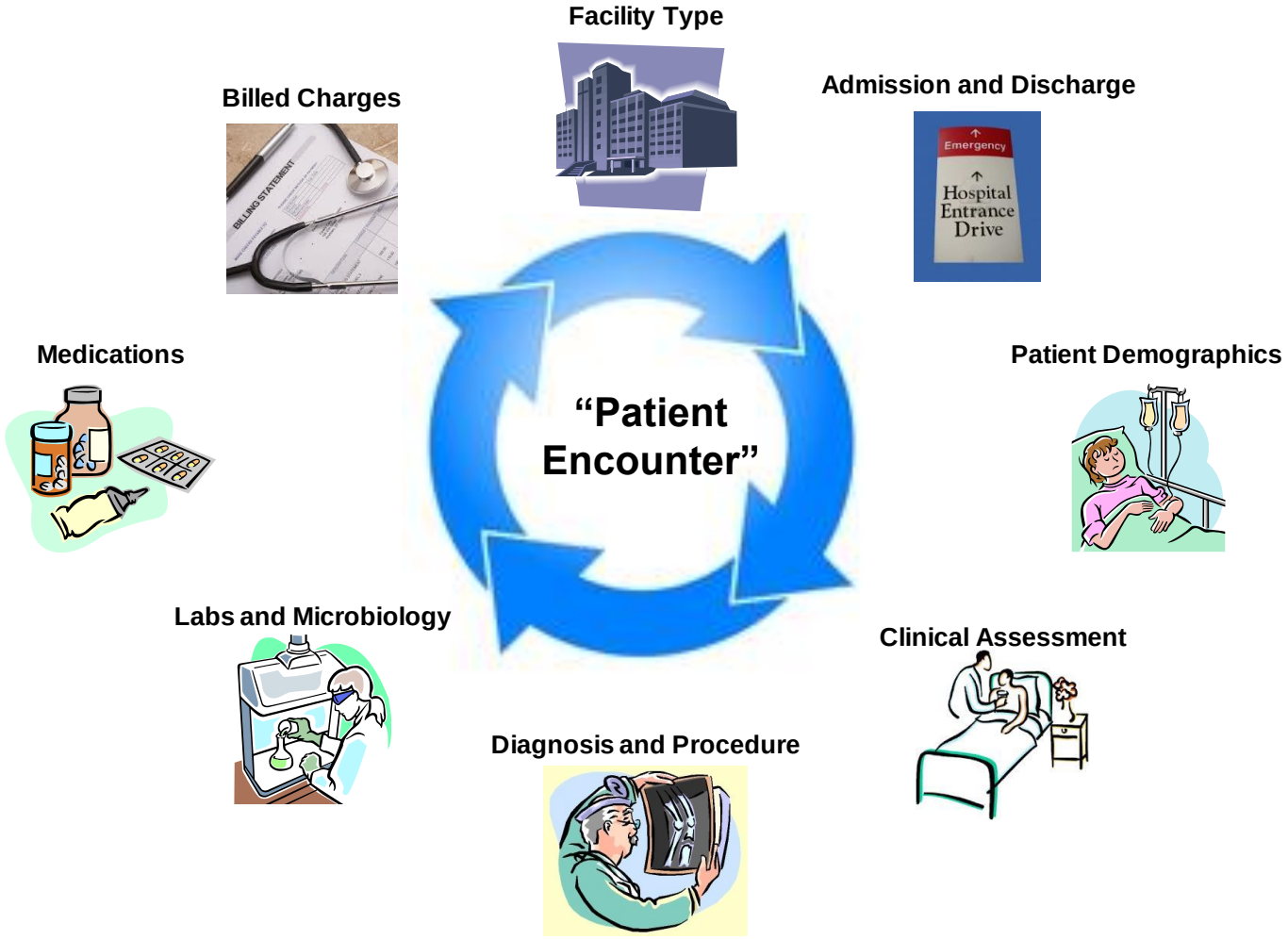
Commercial/ Marketing

- Identify product value attributes
- Forecasting of sales volume and revenue
- Product Competitive positioning
- Product Life cycle management

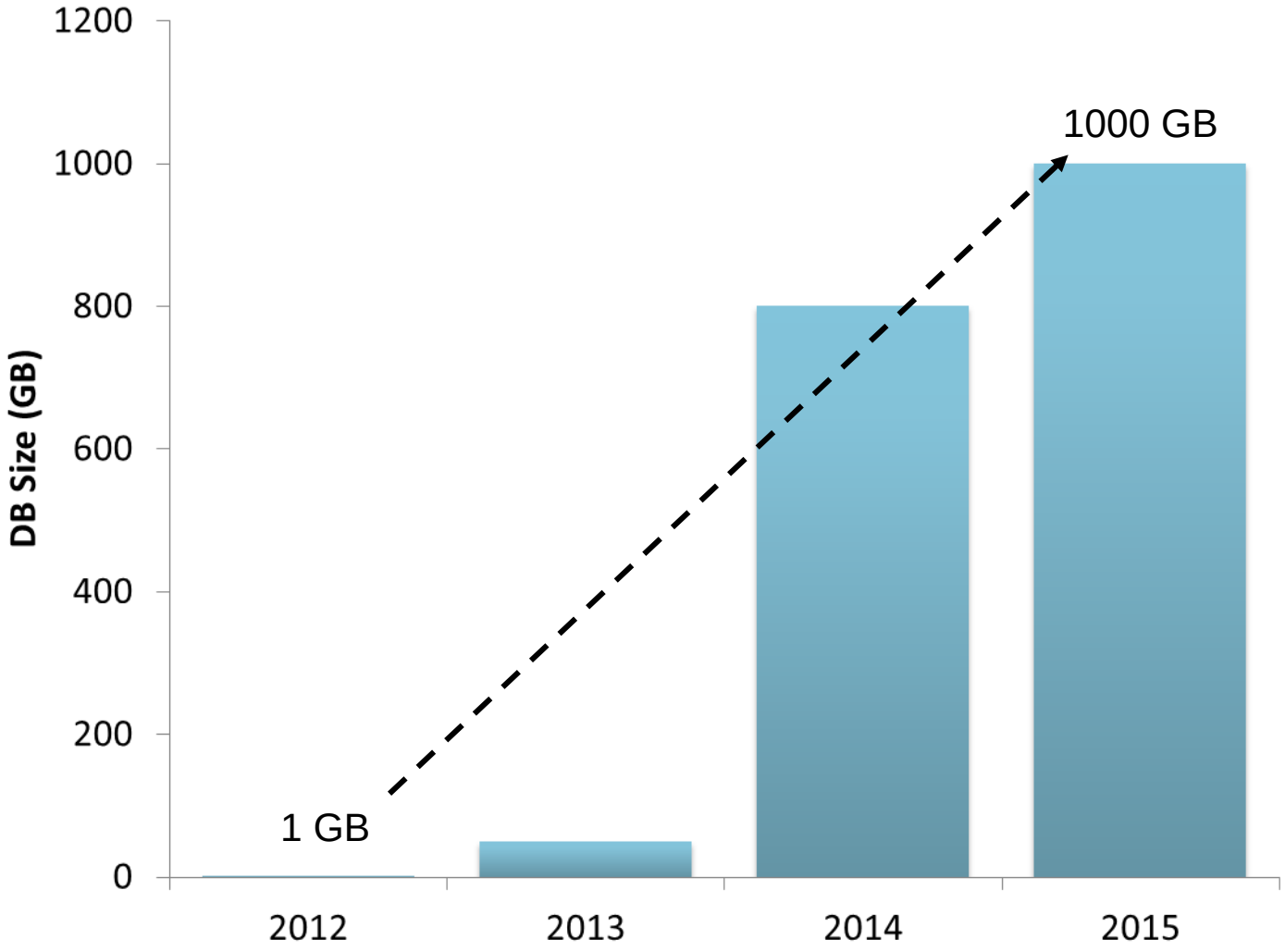
Improved Care Delivery

EHR = Electronic Health Records

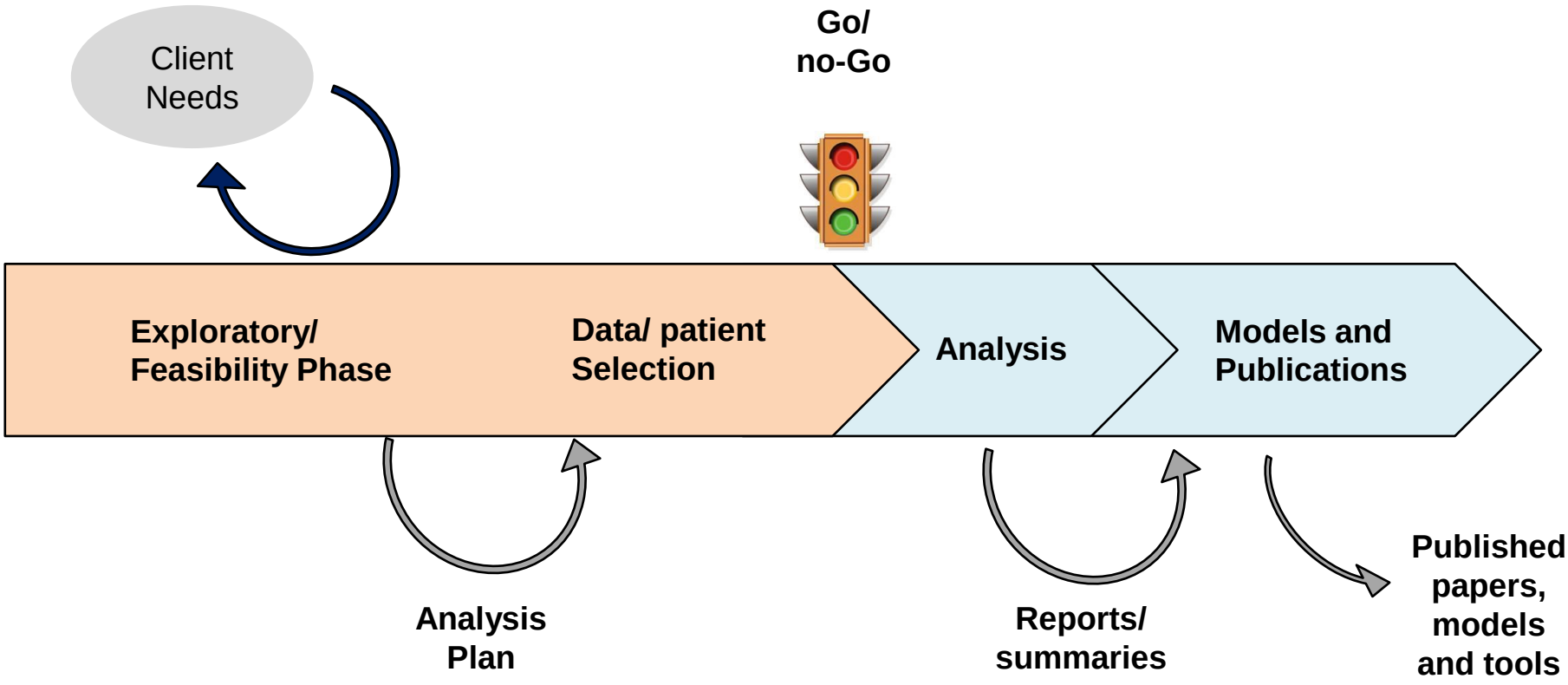
We began working with subsets of Cerner HealthFacts data in 2012 and over time have adopted the much larger, de-identified data set



The jump in data size has opened the door to new opportunities, but has also brought challenges

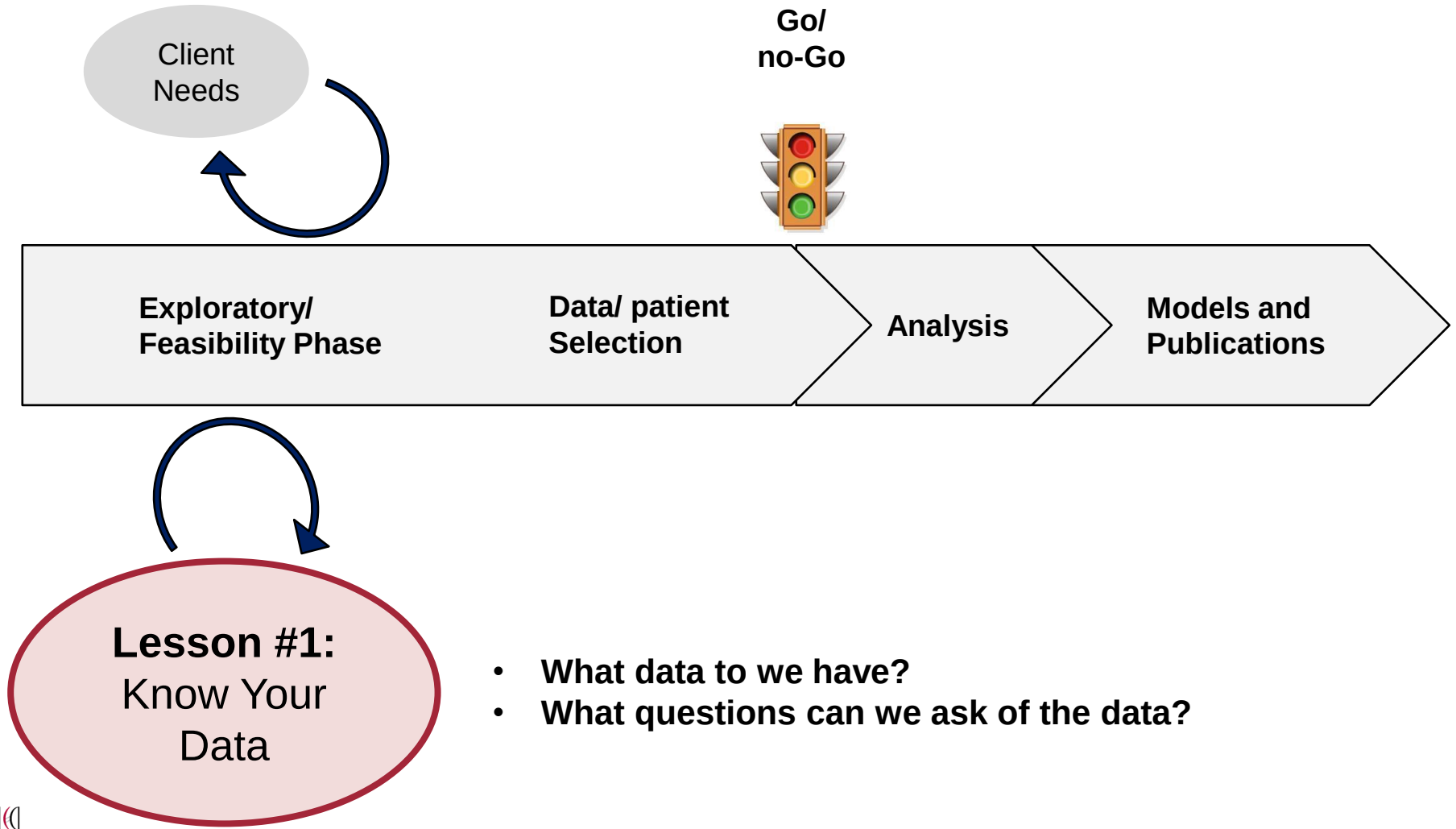


Owing to complexities of real-world data, the exploratory phase determines the feasibility of the study and identifies data gaps and inconsistencies



The iterative nature of the exploratory phase is often the most time consuming aspect of the analysis, compared to other phases

Owing to complexities of real-world data, the exploratory phase determines the feasibility of the study and identifies data gaps and inconsistencies



- **What data to we have?**
- **What questions can we ask of the data?**

Lesson #1: Know Your Data

Plan for discoveries, particularly in the exploratory phase, that move the progress bar back to the beginning and cost many hours of work

Lesson #1: Know Your Data

Approaches to identify dialysis modality (CRRT v. IHD)* in EHR data

Approach	Outcome
Orders for replacement fluid volume (5L bags)	X
Frequency of billed procedure	X
…approach 3	X
…approach 4	X
…approach 5	X
…approach 6	X
…approach 7	X
…approach 8	X
…approach 9	?

8 start-overs, 30 days

*** CRRT = Continuous Renal Replacement Therapy**
IHD = intermittent hemodialysis

Data "dead ends" are difficult to budget for.
Client expectations must be managed appropriately

Seeking advice from data providers can save time and resources

Lesson #1: Know Your Data

Example Approach: To identify renal replacement therapies, we initially looked for medication orders for large-volume fluid orders

Problem: Very few medication orders related to renal replacement therapy

Cause: The bags of fluid are stored on the hospital floor and do not go through the pharmacy, hence they are not captured in the EHR data

Takeaways:
Consulting with our data provider helped inform our team early in the process

Seeking advice from physicians can help understand the workflow on the hospital floor

Lesson #1: Know Your Data

Example Approach: Identify renal replacement modality (continuous vs. intermittent) by analyzing frequency of dialysis procedures

Problem: No consistent pattern in the frequency of dialysis procedures in the EHR data

Cause: Some clinics use punch-cards to keep track of dialysis procedures. Patients could stockpile anywhere from a day to a week of dialysis punch-cards and hand them in all at once

Takeaways:

- This is a deeper problem that was discovered much later in the process -- difficult to plan for
- Our investigator / MD advisor discovered this from a person who works closely with those at the clinic front desk -- upstream in the data gathering process

After several more failed methods, we constructed a more complicated solution

Lesson #1: Know Your Data

Strategy for AKI patient dialysis identification as CRRT v. IHD

Criteria	No. Patient Visits
Hemodialysis [ICD9 procedure code 39.95 ≥1] AND insertion of venous catheter for dialysis[1]	29,005
Diagnosis of AKI[2]	14,645

Excluded n=11,705

CRRT

- Required titrated calcium chloride or calcium gluconate and NOT a bolus injection. (Required an infusion time unit value of 'titration'.) (A) OR required ≥2 bags of calcium chloride or calcium gluconate (B)

N=288 [62 (A) + 226 (B)]

IHD

- Required one of the following conditions:
  - A bolus heparin dose as defined by a dose <10,000 units and no follow-up continuous infusion OR
  - No anticoagulation AND no calcium titration

N=2,652

1) ICD9 procedure code 38.95 Venous catheter for renal dialysis
2) AKI is defined as any one of the following diagnoses (diagnosis of any priority or type):
   - Acute Glomerulonephritis with Other Specified Pathological Lesion in Kidney
   - Acute Kidney Failure with Lesion of Tubular Necrosis
   - Acute Kidney Failure, Unspecified
   - Renal Failure, Unspecified

We successfully created 1:1 and 3:1 propensity matched cohorts

Lesson #1: Know Your Data

Criteria	No. Patient Visits
Hemodialysis [ICD9 procedure code 39.95 ≥1] AND insertion of venous catheter for dialysis[1]	29,005
Diagnosis of AKI[2]	14,645

Excluded n=11,705

CRRT

- Required titrated calcium chloride or calcium gluconate and NOT a bolus injection. (Required an infusion time unit value of 'titration'.) (A) OR required ≥2 bags of calcium chloride or calcium gluconate (B)

N=288 [62 (A) + 226 (B)]

IHD

- Required one of the following conditions:
  - A bolus heparin dose as defined by a dose <10,000 units and no follow-up continuous infusion OR
  - No anticoagulation and no calcium titration

N=2,652

Propensity Match

CRRT
N=185

IHD
N=185

CRRT
N=

IHD
381

Successful but with a caveat…

1) ICD9 procedure code 39.95 Venous catheter for renal dialysis

2) AKI is defined as any one of the following diagnoses (diagnosis of any priority or type); Acute Glomerulonephritis with Other Specified Pathological Lesion in Kidney; Acute Kidney Failure with Lesion of Tubular Necrosis; Acute Kidney Failure, Unspecified; Renal Failure, Unspecified

In matching patients on propensity score, we found critically ill patients being excluded from the matched population

Lesson #1: Know Your Data

Criteria	No. Patient Visits
Hemodialysis [ICD9 procedure code 39.95 ≥1] AND insertion of venous catheter for dialysis[1]	29,005
Diagnosis of AKI[2]	14,645

Excluded n=11,705

CRRT

- Required titrated calcium chloride or calcium gluconate and NOT a bolus injection. (Required an infusion time unit value of 'titration'.) (A) OR required ≥2 bags of calcium chloride or calcium gluconate (B)

N=288 [62 (A) + 226 (B)]

IHD

- Required one of the following conditions:
  - A bolus heparin dose as defined by a dose <10,000 units and no follow-up continuous infusion OR
  - No anticoagulation and no calcium titration

N=2,652

Propensity Match

CRRT
N=

IHD
185

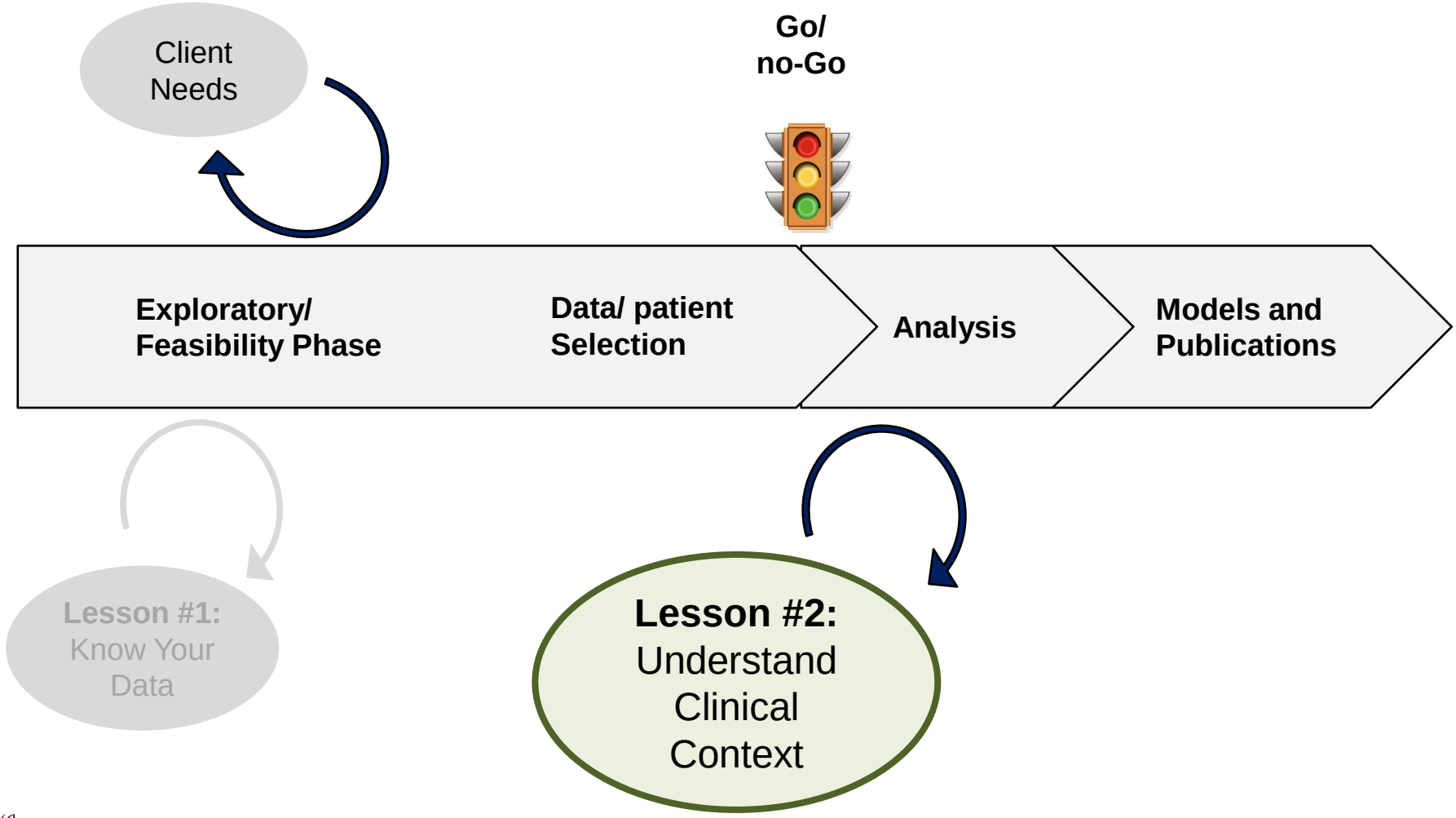
C
N=

D
381

Including severity of illness measure into the propensity score model, excluded Critically Ill patients

1)
2) AKI is defined as any one of the following diagnoses (diagnosis of any priority or type); Acute Glomerulonephritis with Other Specified Pathological Lesion in Kidney; Acute Kidney Failure with Lesion of Tubular Necrosis; Acute Kidney Failure, Unspecified; Renal Failure, Unspecified

It is critical to understand clinical context of an intervention



Lesson #2:
It is critical to understand clinical context of an intervention

In matching patients on propensity score, we found critically ill patients being excluded from the matched population

Lesson #2: Understand Clinical Context

Criteria	No. Patient Visits
Hemodialysis [ICD9 procedure code 39.95 ≥1] AND insertion of venous catheter for dialysis[1]	29,005
Diagnosis of AKI[2]	14,645

Excluded n=11,705

CRRT

- Required titrated calcium chloride or calcium gluconate and NOT a bolus injection. (Required an infusion time unit value of 'titration'.) (A) OR required ≥2 bags of calcium chloride or calcium gluconate (B)

N=288 [62 (A) + 226 (B)]

IHD

- Required one of the following conditions:
  - A bolus heparin dose as defined by a dose <10,000 units and no follow-up continuous infusion OR
  - No anticoagulation and no calcium titration

N=2,652

Propensity Match

CRRT
N=

IHD
185

C
N=

D
381

Including severity of illness measure into the propensity score model, excluded Critically Ill patients

1)
2) AKI is defined as any one of the following diagnoses (diagnosis of any priority or type); Acute Glomerulonephritis with Other Specified Pathological Lesion in Kidney; Acute Kidney Failure with Lesion of Tubular Necrosis; Acute Kidney Failure, Unspecified; Renal Failure, Unspecified

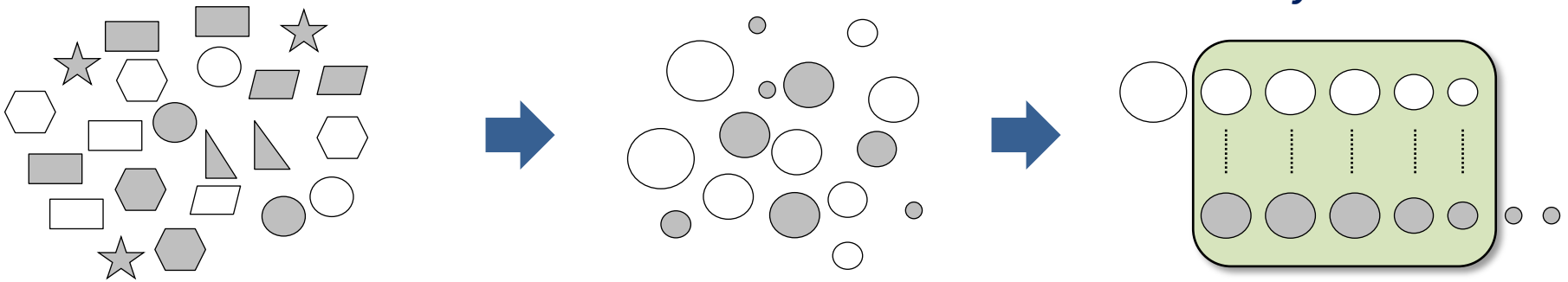
Propensity score methods allow one to minimize the effects of observed confounding when estimating treatment effects using observational data

Lesson #2: Understand Clinical Context

Patients
Varying shapes denote clinical characteristics

Propensity Score Model
Probability of receiving intervention B

Matching on the Propensity Score



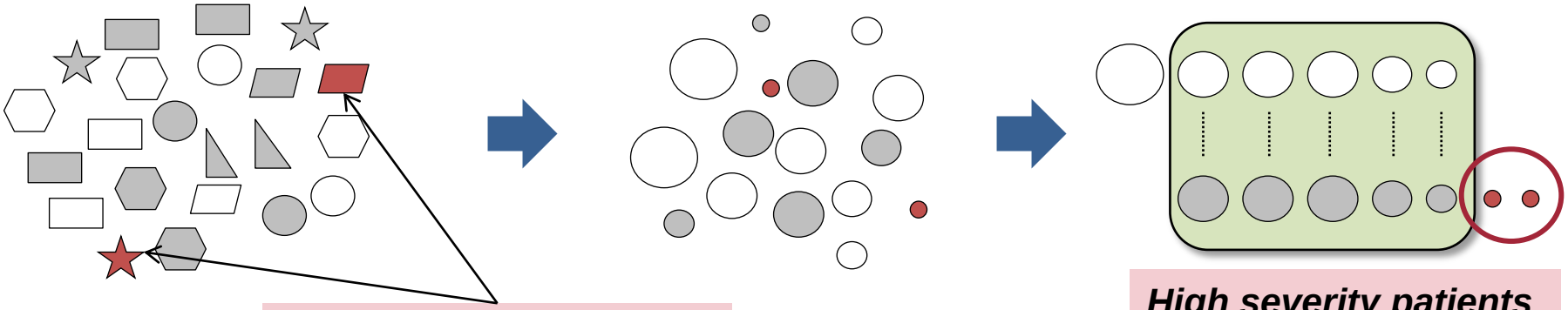
= Patient receiving intervention A

= Patient receiving intervention B

- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. doi:10.1093/biomet/70.1.41
- Figure adopted from: Moss, R. R., Humphries, K. H., Gao, M., Thompson, C. R., Abel, J. G., Fradet, G., & Munt, B. I. (2003). Outcome of mitral valve repair or replacement: a comparison by propensity score analysis. *Circulation*, *108 Suppl* (90101), II90–7. doi:10.1161/01.cir.0000089182.44963.bb

Sicker patients were eliminated during the matching, thus limiting the generalizability of the renal replacement analysis

Lesson #2: Understand Clinical Context



High severity patients were <u>always</u> assigned the same treatment protocol – Continuous renal therapy

High severity patients were excluded by the propensity score matching

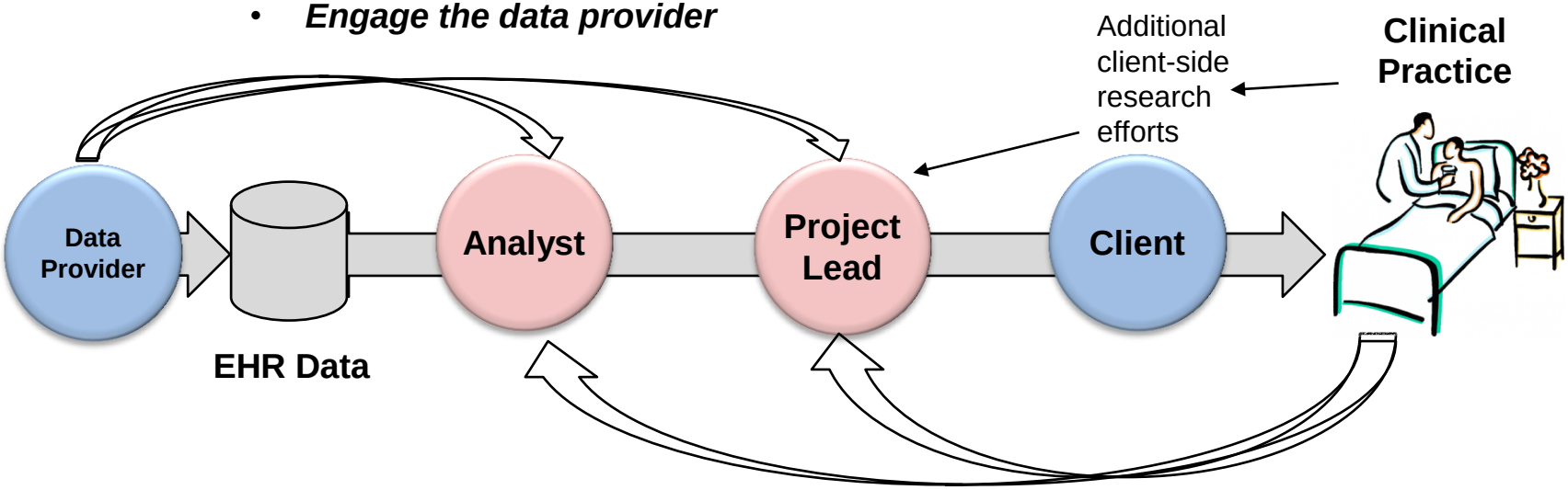
- Both the analyst and clinician needed to be engaged to spot this side-effect
- Recommendation was to conduct a subgroup analysis on the severely ill patients

We found that two types of advisories are especially valuable in uncovering anomalies earlier in the process

Lesson #2: Understand Clinical Context

Seek Data Source Advisory

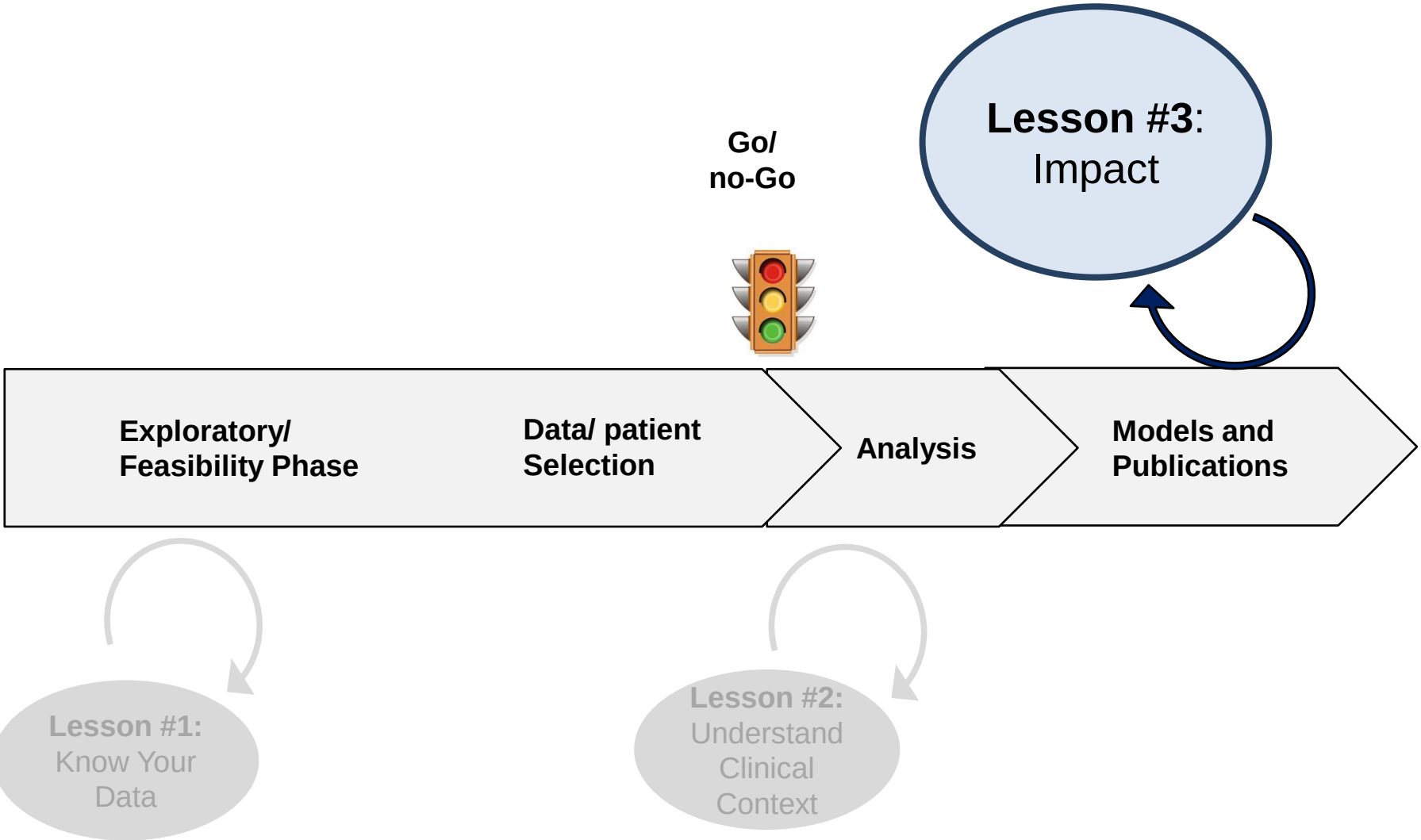
- Issues/pitfalls with underlying data
- Alternative parameters
- Programming considerations
- ***Engage the data provider***



Seek Clinical Practice Advisory

- How practice guidelines may be affecting data
- Exceptions to the rules
- Also reveal patterns in back-office / data entry practices
- ***Engage client to reach out to their experts***
- ***Find our own experts***

Lesson #3: Carry learnings forward and leverage data to develop products beyond a publication



Lesson #3:
Carry learnings forward and leverage data to develop products

Types of impact

Have to know what your impact can be to make sure that your project impactful

Impact can come in many formats

- Research context
  - Seminal paper impacting healthcare practice
- Commercial context
  - Identifying specific patient populations that benefit most from a novel diagnostic
- Health Economics
  - Identifying reduction in costs associated with patient waiting times to save money for the hospital budget and to improve care delivery

Budget impact models present analysis results in actionable format

Lesson #3: Impact

Rationale for developing BIMs:

- With consistent pressure on healthcare entities to cut costs, purchasing a high priced platform needs to be justified in terms of the economic value that it may provide
- Healthcare decision makers are increasingly relying on Budget Impact Models to evaluate the financial impact of adopting new products
- Globally, many healthcare regulatory bodies require financial impact analysis to justify reimbursement

Convert research findings into tools that provide actionable recommendations

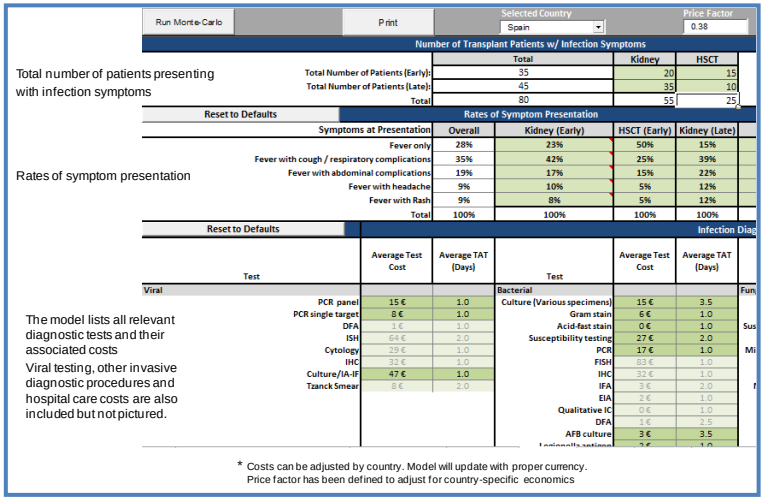
Lesson #3: Impact

Augmenting research findings with assumptions from secondary sources can provide useful insight into the impact of hospital budget

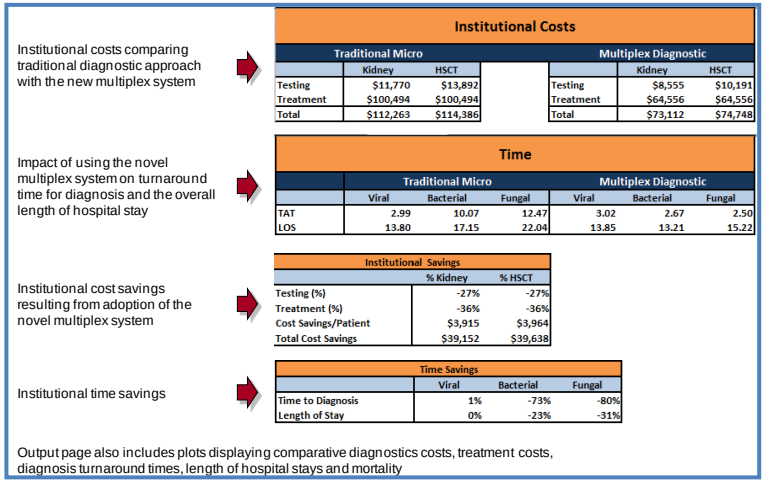
Example findings from renal replacement project:

- Budget model provided evidence to dispel the general belief that the cost center was associated with treatment costs, i.e. cost of CRRT vs IHD
- The largest contributions to cost were downstream procedures
- These two things combined showed the prevailing attitudes were leading to a penny-wise-pound-foolish situation -- this was a surprising find and powerful evidence for the client team to inform their other teams and formulate future strategies

Hospital Specific Inputs



Actionable Outputs



Lesson for a newcomer to “Big Data” analysis:

1. Know your data
    - Patient selection can make or break a project

2. It is critical to understand clinical context of an intervention
    - Consult expert sources, e.g. physicians and nurses

3. Impact
    - Carry learnings forward and leverage data to develop products

Thanks To

Harlen D. Hays, MPH

Fabian D'Souza, MD, MBA, President

Fred Peyerl, PhD, MBA, Partner

David Hayashida, Technology Lead